VISIT…

LANZAROTE Caliente.COM

Современные подходы к интеграции информационных технологий / Труды V Всеросс. семинара "Информационные технологии в энергетике", Иркутск, 2000. Иркутск: ИСЭМ СО РАН, 2001. – 235с. – С.157-165.

НЕЙРОСЕТЕВЫЕ АНАЛИЗ И МОДЕЛИРОВАНИЕ СОВРЕМЕННЫХ СВЯЗЕЙ КЛИМАТА И РАСТИТЕЛЬНОСТИ

В.Г. Царегородцев
Институт вычислительного моделирования СО РАН в г.Красноярске
www.NeuroPro.ru tsar@neuropro.ru
Д.И. Назимова
Институт леса СО РАН в г.Красноярске

Аннотация

Проведена нейросетевая идентификация секторов континентальности и ландшафтных зон Сибири, степени доминирования лесообразователей (трех групп формаций: темнохвойных, сосновых и лиственничных) по климатическим параметрам. В каждой задаче идентификации выделены наиболее значимые климатические параметры. Исследована возможность применения технологии извлечения явных знаний из нейронных сетей применительно к поставленным задачам: в отдельных случаях удалось получить простые линейные границы решения. По сценариям изменения климата на 2030г сделан прогноз возможности смены зонально-секторных характеристик территорий и растительного покрова. Обсуждается возможность применения искусственных нейронных сетей для решения разнообразных задач прогнозирования и классификации при моделировании поведения природных систем.

1. Введение

Глобальные изменения природной среды и климата, вызванные антропогенными и иными причинами, делают актуальным прогноз грядущих трансформаций растительного покрова [1]. В [2-5] отмечается, что ожидаемый сценарий будущего климата сопоставим с таковым в климатическом оптимуме голоцена (около 5,5–6 тыс. лет назад), а следовательно, можно проводить некоторые параллели с реконструкциями ландшафтов и экосистем на этот период истории планеты. Такой подход предполагает взвешенный анализ современных и прошлых связей между климатом и другими компонентами зональных экосистем.

В последние десятилетия достигнуты заметные успехи в исследовании истории ландшафтов Сибири [6]. Хотя развитие покрова протекало по законам эволюции и было, в принципе, необратимым, современные зоны растительного покрова на территории Сибири в наибольшей степени определяются историей климата. То же можно сказать и о почвенном покрове, хотя он и хранит заметные черты экосистем прошлого. Поэтому задача идентификации связей между климатом и покровом является значимой и для палеогеографии, и для прогноза. Имеется удачный опыт моделирования биомов суши на основе связей между параметрами климата и зональными классами покрова [4].

К числу зональных классов покрова Сибири, рассматриваемых в работе, мы относим биомы и зоны. Биом [7] представляет единство зонального (регионального) климата с биотой и субстратом. Перечень биомов отражен в биомной модели глобального уровня [4] и в моделях растительности Сибири, опубликованных в 90-е годы.

Цель работы – опираясь на разработанные биоклиматические модели, провести исследование возможностей применения технологии нейронных сетей для прогноза смен зон и биомов Сибири в меняющемся климате, включая исследование подмножеств климатических признаков с точки зрения их значимости для решения данной задачи.

Поскольку имеются различные экспертные сценарии будущего изменения

климата, исследование влияния отдельных климатических параметров на состояние покрова поможет в дальнейшем проводить уточнение и адаптацию сценариев.

2. Исходные данные

Информационной основой послужила база данных, включающая характеристики зональных типов растительности: состав зональных формаций и групп типов леса, структуры почвенного покрова, уровень продуктивности, количественное соотношение фоновых экобиоморф в зональных условиях и т.п. Данные систематизированы по территориям, с привязкой к метеостанциям. Территории охарактеризованы по данным метеостанций следующими средними многолетними климатическими показателями: суммами осадков за год, теплый и холодный периоды года, средними температурами января и июля и годовой суммой температур, годовой амплитудой температур, годовыми суммами температур выше 0°С и выше 10°С, индексом континентальности Конрада [8], коэффициентом увлажнения Мезенцева – всего 12 показателей, в интегральной форме отражающих тепло- и влагообеспеченность регионального климата. Хотя выборка и не характеризует всего разнообразия локальных климатов, тем не менее, позволяет увидеть пространственно упорядоченную картину изменения биоклиматических характеристик от Урала до Дальнего Востока. Общее число станций в базе данных для территории Сибири более 500, из которых использовались 416 с наиболее полной информацией. Ранее [9-11] использовался несколько отличающийся набор климатических показателей при меньшем объеме базы данных.

Положение точки (метеостанции) в пространстве климатических параметров более однозначно определяет все зональные характеристики лесных массивов, чем на положение на карте. Точность географического прогноза современной растительности по параметрам климата достаточно велика: это показала проверка рабочих качеств модели растительности Сибири путем сравнения с ландшафтной картой региона [12]. Привлекательным для использования в целях прогноза является и то, что само положение точки относительно осей климатических координат дает много разнообразной информации о свойствах объекта, действительно определяемых климатом. К их числу можно отнести не только признаки почвенной и лесорастительной структуры и состава, но и признаки функционирования экосистем – сезонную ритмику, уровень продуктивности и т.п. Т.к. климат – система многомерная и разномасштабная (макро-, мезо-, и микроклиматы), предстоит выяснить, какие подмножества признаков климата, имеющихся в базе данных, наиболее информативны для решения задач прогноза и диагностики зональных классов экосистем.

3. Используемые методы

Особенностью рассматриваемой базы данных является то, что зонно-секторное деление территорий выполнялось экспертами на основе почвенно-растительных признаков. Целью же является воспроизведение экспертной классификации на основе признаков климата. При этом ни в исходном пространстве почвенно-растительных признаков, ни в пространстве климатических параметров не наблюдается четких границ между классами – возможно глубокое проникновение объектов одного класса в множество объектов другого класса. Также в признаковом пространстве не существует и кластерной структуры, хорошо соответствующей зонно-секторному делению [13].

Поэтому использование классических методов построения линейных разделяющих поверхностей в задаче классификации наталкивается на невозможность на настоящей базе данных получить правильную классификацию всех объектов обучающей выборки.

Искусственные обучаемые нейронные сети [14,15] позволяют строить сильно нелинейные разделяющие поверхности, при этом число классов в классификационной модели может быть произвольным, не требуется задавать класса функций, к которому

должна принадлежать разделяющая поверхность. Имеется возможность работы как с непрерывными, так и с дискретнозначными и номинальными зависимыми и независимыми признаками; число примеров обучающей выборки для разных классов может быть различным. Имеется процедура определения значимости независимых признаков и сокращения их числа. В последнее время разрабатывается технология извлечения явных знаний из нейросети – способ записи сформированного нейросетью навыка решения неформализованной задачи предсказания или классификации в виде, доступном для последующей семантической интерпретации пользователем [16,17].

Использовалась авторская программа-нейроимитатор NeuroPro, позволяющая проводить анализ и моделирование данных с использованием нейросетевых и классических статистических методов, исследовать и визуализировать исходные данные и результаты моделирования, выполнять извлечение явных знаний из нейронных сетей.

4. Зонально-секторная идентификация территорий

Рассматривались секторы континентальности – континентальный, резко континентальный, крайне континентальный, а также ландшафтные зоны – тундра и лесотундра, тайга и подтайга, лесостепь, степь. Хотя у секторов континентальности и есть некоторая аналогия с западно-, средне- и восточносибирским ландшафтно-климатическими секторами, эта аналогия не является полной, так как секторы климатического пространства не совпадают с географическими. В одном географическом секторе могут быть совмещены два сектора континентальности и два-три сектора увлажнения, что характерно для горных территорий. Это вызывает сочетание сразу нескольких биомов в одном горном ландшафте.

Выделились следующие задачи идентификации:

1. Идентификация границ ландшафтных зон, выполняемая путем построения бинарных классификаций "тундра+лесотундра – тайга+подтайга", "тайга+подтайга – лесостепь", "лесостепь – степь".
2. Идентификация границ секторов континентальности для равнинных метеостанций, выполняемая как путем построения классификации на 3 класса (для всех трех секторов континентальности одновременно), так и через попарное разделение соседних секторов континентальности.

Напомним, что начальное определение зонально-секторной принадлежности осуществлялось экспертами на основе лесорастительного районирования и тематических карт. Нейросети же обучались идентифицировать секторно-зональную принадлежность только по признакам климата. Эксперименты показали высокую точность идентификации – проверка путем скользящего экзамена давала 77-93% правильных ответов для разных постановок задач классификации. Полученные уровни точности решения во многом легко и однозначно интерпретируются: например, зона лесостепи на территории Сибири, наиболее освоенная человеком и наиболее подвергшаяся антропогенным нарушениям, только по признакам климата идентифицируется плохо, как и можно было ожидать.

Задача определения значимости климатических признаков решалась как задача нахождения минимального набора признаков, на основании которого нейросеть безошибочно решает задачу. Для каждой классификационной задачи создавалось по 30 нейросетей, далее для каждой нейросети проводилось сокращение числа признаков путем последовательного нахождения наименее значимого признака, исключения его из числа входных сигналов нейросети и дообучения нейросети. Затем выполнялся анализ полученных минимальных наборов по каждой классификации. Значимостью признака считалась частота появления признака в полученных минимальных наборах.

Для всех классификационных задач удалось найти 2-5 наиболее значимых признаков из исходных 12-ти. При этом повторное решение задач идентификации на основе только наиболее значимых признаков не приводило к снижению точности по

сравнению с достигнутой на 12-ти признаках точности идентификации. Это говорит о том, что исходный набор признаков избыточен для решения поставленных задач.

Экспертный анализ выявленных наиболее значимых признаков показал, что для большинства случаев значимые признаки не совпали с традиционно используемыми экспертными оценками [9]. Однако, в некоторых случаях удавалось найти правдоподобное объяснение причин значимости выделенных нейросетями признаков.

С целью выявления тенденций изменения значимости признаков при изменении географической долготы были решены несколько задач попарного разделения соседних ландшафтных зон внутри одного сектора континентальности. К сожалению, не для всех случаев удавалось сформировать представительную обучающую выборку. Но все же удалось установить, что значимость признаков при решении некоторой задачи разделения ландшафтных зон может меняться при переходе от сектора к сектору, и что значимости внутри сектора могут отличаться от значимостей для той же задачи разделения зон для всех трех секторов континентальности.

Удалось подтвердить известный результат, что с изменением географических долготы и широты меняется и вид границ между различными классами покрова: одну из важнейших черт восточносибирских и части южносибирских ландшафтов составляет глубокое проникновение фрагментов степного биома в таежную зону, а тундрового – в горно-таежную, а в Западной Сибири хорошо выражена пограничная зона мелколиственной подтайги на переходе от тайги к лесостепи. Так, для Западной Сибири удалось провести границу между классами "лес" и "степь" по единственному параметру – либо по годовой сумм температур выше 0°С, либо по годовой сумме температур выше 10°С, либо по средней температуре июля, а в то же время для Восточной Сибири не удалось выявить одного-двух признаков, по которым наблюдается четкое отделение этих двух классов друг от друга (надо заметить, что здесь при разделении лесных и степных территорий специально был опущен промежуточный лесостепной класс).

Для каждой классификационной модели после выявления наиболее значимых признаков делалась попытка записи границ решения (разделяющей поверхности между классами) в наиболее простом виде в пространстве выделенных наиболее значимых признаков. Для этого применялась технология извлечения явных знаний из нейронной сети [16,17] путем максимального упрощения нейросети и последующей записи набора продукционных правил логического вывода по полученной нейронной сети. В отдельных случаях удавалось построить линейную разделяющую поверхность в пространстве одного-двух признаков (точку или прямую), отделяющую объекты одного класса от объектов другого класса. При этом число объектов, ошибочно относимых к другому классу, оказывалось малым (до 5% неправильно идентифицируемых объектов в выборке), и экспертный анализ этих объектов с привлечением дополнительной информации (о почвенно-гидрологическом режиме покрова, антропогенных нарушениях,) показывал, что такие объекты действительно более похожи на объекты другого класса.

5. Исследование возможности смены секторно-зональной принадлежности при заданном сценарии изменения климата

Основываясь на тех климатических параметрах, прогноз изменений которых на будущее наиболее прост и наименее противоречив, в [9] исследовалась возможность смены секторной и зональной принадлежности станций при заданном сценарии изменения климата.

В качестве параметров были взяты температуры января и июля, по которым затем рассчитывается индекс континентальности Конрада. Также взяты суммы осадков за год и холодный период года, хотя прогнозируемые изменения величин осадков неточны и противоречивы [4,18].

На выборке с пятью указанными показателями нейросети были обучены решению

ранее описанных задач зонно-секторной идентификации. Несмотря на уменьшение числа показателей, снижения точности решения задач не произошло.

Прогноз изменения климата был сделан с учетом усредненных экспертных данных климатологов на 2030г: увеличение температур июля на 1-2°С, января – на 3-4°С, годовых сумм осадков – на 30-50мм, сумм осадков за холодный период года – на 10-30мм. Были взяты 40 метеостанций (расположенные в разных ландшафтных зонах и секторах континентальности), и их климатические параметры были соответственно изменены с учетом местоположения так, что максимальные изменения по условию прогноза задавались для территорий севера и востока. Затем с помощью обученных нейросетей определялась новая (потенциальная) зонально-секторная принадлежность станций на основе данных, откорректированных в соответствие с прогнозом.

Результаты показали, что на территории Сибири есть точки (регионы), которые при смене климата могут изменить секторную и зональную принадлежность [9].

При заданном изменении температурного режима и небольшом увеличении норм осадков многие географические точки приобретают биоклиматический потенциал, свойственный соседней, более южной зоне, и/или соседнему, как правило менее континентальному сектору. Соответственно, можно ожидать тенденций смены состава типов растительности и видов лесообразователей, усиления конкурентоспособности одних пород, ослабления других. Однако, понятно, что релаксация лесорастительного покрова к новому равновесному состоянию, соответствующему измененному климату, будет требовать достаточно длительного времени (десяти-сотни лет).

Экспертный анализ показал, что для многих случаев компьютерный прогноз более обоснован, чем предположения отдельных участников экспертизы – благодаря многомерности модели. Для 26 метеостанций секторная и зональная принадлежность остались прежними. Для 8 станций изменилась зональная принадлежность, для 7 станций изменился сектор континентальности. Чаще всего прогноз выглядит правдоподобным. И все же специалисту иногда трудно согласиться с вероятностью тенденций, указанных нейросетями. Так, тенденция к смене лиственничников сосняками в Якутии оценивается экспертами как пока нереальная – сказывается недоучет почвенного гидротермического режима вечной мерзлоты – мощного консервативного фактора.

При указанном потеплении, даже при увеличении годовых сумм осадков на 30-50 мм (что выглядит реалистичным, с учетом палеогеографических реконструкций), увлажнения окажется недостаточно, чтобы поддержать современное состояние лесистости на южной окраине бореальной области: лес здесь, скорее всего, уступит место лесостепи. Только увеличение годового количества осадков на 200 и более мм при росте мощности снежного покрова позволит сохраниться лесным формациям. Можно ожидать изменения лесорастительного потенциала и во многих других зональных типах ландшафтов, в частности, усиления позиций лиственных пород на месте темнохвойных. Этот прогноз делается на основе связей хвойных и лиственных формаций с климатом.

Выполненное в [9] моделирование изменения климата опиралось на выбранные экспертами пять климатических параметров, однако, эти параметры для большинства рассмотренных задач классификации не являются наиболее информативными (см. Раздел 4). Возникает идея использования для каждой задачи классификации своего набора информативных параметров и моделирования изменения именно их. Естественно, здесь возникают трудности при оценке величины будущего изменения выделенных параметров, но зато появляется более адекватное представление о сущности протекающих процессов.

6. Определение значимости климатических параметров для идентификации зональных групп лесообразователей

При прогнозировании смены зонально-секторной принадлежности территории

сделать однозначный качественный (по составу) или количественный (по степени доминирования) вывод о смене лесообразователей на данной территории нельзя. Для таких прогнозов нужно вводить самостоятельные прогнозные модели.

Попытка определения степеней доминирования лесообразователя по признакам климата дала плохие точности решения, в основном из-за слишком неравномерной и/или недостаточной представительности обучающих примеров отдельных классов. Поэтому пришлось перейти к отдельным более грубым постановкам – бинарной классификации лесных массивов по преобладанию того или иного типа лесообразователей. Попарно отделялись друг от друга следующие классы по составу лесообразователей – темнохвойные (кедр, пихта), лиственничные и сосновые. Это дало достаточную точность идентификации и позволило перейти к определению значимости параметров [10,11].

Наиболее значимыми для отделения зональных классов темнохвойных лесов от лиственничных оказались индекс континентальности, сумма осадков в холодное время года, годовая сумма осадков. Параметры теплообеспеченности значимыми не были. Эти результаты находятся в полном согласии с известными закономерностями географии кедра, ели, пихты и лиственницы, с широкими диапазонами их возможного произрастания по градиенту тепла и с четкой границей по индексу континентальности.

При разделении сосновых и лиственничных зональных массивов наиболее значимыми оказались коэффициент влажности, осадки за холодного периода года и годовые осадки. При отделении темнохвойных зональных массивов от сосновых четкого выделения наиболее значимых факторов не получилось, но все параметры теплоообеспеченности оказались незначимыми. Видимо, в последнем случае роль эдафических факторов перекрывает по значимости климатические.

Знания об информативности признаков помогут интерпретировать результаты моделирования потенциальной смены состава лесообразователей на территории при изменении климата.

7. Заключение

Проведенные эксперименты и достигнутая точность идентификации подтвердили вывод ранее выполненной работы [9] о возможности идентификации с достаточной точностью крупных ландшафтных категорий по нескольким климатическим показателям. Зоны тайги, подтайги, черневых лесов, традиционно выделяемые по признакам растительности и почвенного покрова, также могут быть идентифицированы с высокой степенью надежности только по признакам климата.

Исследования показали, что нейронные сети являются мощным инструментарием для решения задач классификации и прогноза ландшафтных категорий по признакам климата. Прогнозирование может быть полезным для выявления экстремальных ситуаций для определенных природных и антропогенных экосистем, например, для темнохвойных лесов на их пределах, определяемых дефицитом увлажнения климата. Те лесообразователи, которые в новой климатической ситуации вышли за пороговые значения их устойчивости, оказываются в зоне повышенного риска и даже могут элиминировать. Это новый аспект анализа вдобавок к традиционному использованию знаний экологии видов-лесообразователей, их конкурентных взаимоотношений, устойчивости к пожарам и инвазиям насекомых-вредителей, и т.д.

Экологическая интерпретация биогеографических закономерностей не может быть дана только на основании современных связей между растительным покровом и климатом. Необходим учет и палеогеографии, и влияния совокупности других факторов, осложняющих проявление биоклиматических зависимостей. На территории Сибири к ним относится прежде всего мерзлота почвогрунтов. Существуют и другие причины, например, обратные связи лес–климат, специфика почвенно-гидрологического режима ландшафтов, наконец, широкомасштабные антропогенные нарушения, препятствующие

однозначному проявлению биоклиматических закономерностей. Поэтому немаловажно выявить реальную меру упорядоченности зональных признаков лесного покрова в многомерном пространстве климатических и иных показателей.

Таким образом, просматривается перспективное направление нейросетевого моделирования поведения природных систем при глобальных изменениях климата.

ЛИТЕРАТУРА

1. Устойчивое развитие бореальных лесов // Тр. VII ежегодной конференции МАИБЛ. – С.Петербург, 1997.
2. Будыко М.И. Климат в прошлом и будущем. – Л., 1980.
3. Величко А.А. Зональные и макрорегиональные изменения климатических условий, вызванные "парниковым эффектом" // Изв. АН РАН. Сер. географ. – 1992, №2.
4. Prentice I.C., Cramer W., Harrison S. et all. A global biome model predicting global vegetation patterns from plant physiology and dominance, soil properties and climate // J. Biogeogr. – 1992, v.19.
5. Pernetta J., Leemans R., Elder D. et al. The impact of climate change on ecosystems and species: Terrestrial ecosystems // IUCN. – Gland, 1995.
6. Проблемы реконструкции климата и природной среды голоцена и плейстоцена Сибири. – Новосибирск, 1998.
7. Walter H., Box E. Global classification of natural terrestrial ecosystems. // Vegetatio. – 1976, v.32.2.
8. Conrad V. Usual formulas of continentality and their limits of validity // Trans. of American Geophysical Union. – 1948, v.27(5).
9. Царегородцев В.Г., Погребная Н.А. Нейросетевые методы обработки информации в задачах прогноза климатических характеристик и лесорастительных свойств ландшафтных зон // Методы нейроинформатики. Красноярск: КГТУ, 1998. - 205с. – С.65-110.
10. Назимова Д.И., Царегородцев В.Г. Нейросетевая идентификация зональных групп лесных формаций Сибири // Всеросс. научно-техн. конф. "Нейроинформатика-2000". Сборник научных трудов. В 2 частях. Ч.2. - М.: МИФИ. 2000. - 236с. – С112-119.
11. Царегородцев В.Г., Назимова Д.И. Нейросетевое исследование значимости климатических параметров для идентификации зональных групп лесообразователей Сибири // Тез. докл. VII Всеросс. семинара "Нейроинформатика и ее приложения" / Красноярск: КГТУ. 1999. - 167с. – С.152-153.
12. Tchebakova N.M., Monserud R.A., Nazimova D.I. A Siberian vegetation model based on climatic parameters // Canad. J. of Forest Research. 1995, v.4. – pp.1595-1607.
13. Назимова Д.И., Гуревич М.Ю., Кофман Г.Б. и др. Опыт многомерной классификации лесорастительных районов Сибири по климатическим признакам // География и природные ресурсы. 1997, №2.
14. Горбань А.Н. Обучение нейронных сетей. М.: ПараГраф, 1990. – 160с.
15. Горбань А.Н., Россиев Д.А. Нейронные сети на персональном компьютере. – Новосибирск: Наука, 1996. – 276с.
16. Царегородцев В.Г. Производство полуэмпирических знаний из таблиц данных с помощью обучаемых искусственных нейронных сетей // Методы нейроинформатики. Красноярск: КГТУ, 1998. - 205с. – С.176-198.
17. Царегородцев В.Г. Извлечение явных знаний из таблиц данных при помощи обучаемых и упрощаемых искусственных нейронных сетей // Проблемы нейрокибернетики (Материалы XII Межд. Конф. по нейрокибернетике. Ростов-на-Дону, 1999). - Ростов-на-Дону. Изд-во СКНЦ ВШ. 1999. - 323с. – С.245-249.
18. Tchebakova N.M., Monserud R.A., Leemans R., Nazimova D.I. Possible vegetation shifts in Siberia under climatic change // Impacts of climate change on ycosystems. 1995. – pp.67-83.